Project Overview

Intro to Bayesian statistics has a course project worth 30% of the final grade. The project will be completed individually on a dataset of the student’s choosing. The project should follow the five steps of a Bayesian analysis fro DBDA:

1. Identify the data relevant to the research question(s).
2. Define a descriptive model for the relevant data.
3. Specify a prior distribution on the parameters.
4. Use Bayesian inference to re-allocate probability across parameter values.
5. Check that the posterior predictions mimic the data with reasonable accuracy.

Furthermore, while this class is a Bayesian statistics course, it is also a course on statistical modeling in general. When conducting analyses remember the following QQQ framework:

• **Qualitative**: Define and describe the research question without using statistical lingo in the language of the specified domain. In a collaborative setting, this step is typically done with a collaborator with expertise in a scientific area.

• **Quantitative**: Perform the analysis - 5 steps listed above.

• **Qualitative**: Translate the statistical results (posterior in this case) and make inferences in the language of the specified domain.

All written documents will be completed through R Markdown to enable reproducibility.

Evaluation

A full rubric will be provided later, but the evaluation will be based on:

- Project Proposal: 10%
- Intermediate Project Summary: 10%
- Oral Project Presentation: 30%
- Written Project Description: 40%
- Peer Feedback: 10%

Checkpoints

- March 8: Project Proposal Due
- March 22: Project Proposal Approved
- April 12: Intermediate Project Summary Due
- April 30: Project Due
  - ~6 minute oral description during scheduled final exam period
  - written report with reproducible code due

Project Proposal

The project proposal will contain three parts:

1. A written description of the research question.
2. A description of the dataset you propose using for the project (along with a link).
3. A few paragraphs describing the statistical model you will use to answer your research question.
**Intermediate Project Summary**

The intermediate project summary will build upon the project proposal and have the following sections:

1. An introductory section that describes your study and why it is important.
2. A description of the dataset you propose using for the project along with a few exploratory graphics about the dataset.
3. A detailed section describing the statistical model you will be using, along with clearly defined notation.
4. A section describing the priors you have selected for the parameters in your model and a clear justification of why they are appropriate.

**Written Project Description**

The written project will be the final report that continues to build upon the intermediate project summary. The report should have the following sections (or similar):

1. Introduction
2. Data
3. Statistical Model
4. Priors
5. Computation - include enough written details to describe how you are fitting models, but include all code in the appendix
6. Results - Discuss convergence of your MCMC, and results of the statistical models
7. Discussion - this is the last qualitative section, translate statistical results to context of the problem.

**Oral Project Presentation**

The oral projects will be during the scheduled final exam period on April 30th from 2:00 - 3:50 PM. Each student will have approximately 6 minutes to describe their work. Note: this will not be enough time to discuss the project in great detail, but should give classmates a sense of what was done with the project and what the findings were.

**Peer Feedback**

Throughout the course of the project you will be asked to evaluate classmates writing and presentations. Thoughtful, and respectful, comments will be expected as part of this component.